2017

Mapping for the future: Business intelligence tool to map regional housing stock

Murad Safadi
University of Wollongong, murad@uow.edu.au

Jun Ma
University of Wollongong, jma@uow.edu.au

Rohan Wickramasuriya
University of Wollongong, rohan@uow.edu.au

Daniel J. Daly
University of Wollongong, ddaly@uow.edu.au

Pascal Perez
University of Wollongong, pascal@uow.edu.au

See next page for additional authors

Mapping for the future: Business intelligence tool to map regional housing stock

Abstract

The amount of data available and the lack of data integration represent an increasing challenge to effective planning for government agencies. Integration of data from multiple sources has the potential to enable a user to draw valuable insights, which can be used to enhance service targeting and delivery, and to improve program evaluation. In recognition of the need to improve data integration the University of Wollongong and the NSW Office of Environment and Heritage (OEH) partnered to create an integrated housing stock database for the Illawarra region. The database serves as the backbone for an online and interactive Housing Stock Mapping Dashboard (HSMD). It assembled multilevel granular information (including at the Statistical Area Level 1 (SA1) and Local Government Area (LGA) level) collected from multiple historical programs by multiple agencies. This centralised, integrated data repository can help agencies understand the existing housing stock, and improve access to information to support evidence-based policy. This paper presents a model of how data can be integrated from multiple agencies to provide an online collaboration platform. The platform, HSMD, was designed to demonstrate to government, industry, and the research community the opportunity of data integration and advanced analytics. Potential applications of the HSMD include characterisation of the existing housing stock according to a range of building attributes, for instance the presence of ceiling insulation or rainwater tanks. Comparison of these attributes with energy consumption data can indicate the influence of the attribute, or the impact of a specific intervention. This can help policy makers understand uptake and penetration of previous rebate schemes.

Keywords

intelligence, business, future:, mapping, housing, regional, map, stock, tool

Disciplines

Engineering | Physical Sciences and Mathematics

Publication Details

Safadi, M., Ma, J., Wickramasuriya, R., Daly, D., Perez, P. & Kokogiannakis, G. (2017). Mapping for the future: Business intelligence tool to map regional housing stock. Procedia Engineering, 180 1684-1694.

Authors

Murad Safadi, Jun Ma, Rohan Wickramasuriya, Daniel J. Daly, Pascal Perez, and Georgios Kokogiannakis

Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 180 (2017) 1684 – 1694

Procedia Engineering

www.elsevier.com/locate/procedia

International High- Performance Built Environment Conference – A Sustainable Built Environment Conference 2016 Series (SBE16), iHBE 2016

Mapping for the future: Business intelligence tool to map regional housing stock

Murad Safadi[a,*], Jun Ma[a], Rohan Wickramasuriya[a], Daniel Daly[b], Pascal Perez[a], Georgios Kokogiannakis[b]

[a] *SMART Infrastructure Facility, University of Wollongong,Northfields Avenue, Wollongong, NSW, 2522, Australia*
[b] *Sustainable Buildings Research Centre, University of Wollongong,Squires Way, Wollongong, NSW, 2519, Australia*

Abstract

The amount of data available and the lack of data integration represent an increasing challenge to effective planning for government agencies. Integration of data from multiple sources has the potential to enable a user to draw valuable insights, which can be used to enhance service targeting and delivery, and to improve program evaluation. In recognition of the need to improve data integration the University of Wollongong and the NSW Office of Environment and Heritage (OEH) partnered to create an integrated housing stock database for the Illawarra region. The database serves as the backbone for an online and interactive Housing Stock Mapping Dashboard (HSMD). It assembled multilevel granular information (including at the Statistical Area Level 1 (SA1) and Local Government Area (LGA) level) collected from multiple historical programs by multiple agencies. This centralised, integrated data repository can help agencies understand the existing housing stock, and improve access to information to support evidence-based policy.

This paper presents a model of how data can be integrated from multiple agencies to provide an online collaboration platform. The platform, HSMD, was designed to demonstrate to government, industry, and the research community the opportunity of data integration and advanced analytics. Potential applications of the HSMD include characterisation of the existing housing stock according to a range of building attributes, for instance the presence of ceiling insulation or rainwater tanks. Comparison of these attributes with energy consumption data can indicate the influence of the attribute, or the impact of a specific intervention. This can help policy makers understand uptake and penetration of previous rebate schemes.

Peer-review under responsibility of the organizing committee iHBE 2016

Keywords: Energy epidemiology; housing stock mapping; energy; visualisation; built environment.

* Corresponding Author.

E-mail Address: murad@uow.edu.au

Peer-review under responsibility of the organizing committee iHBE 2016
doi:10.1016/j.proeng.2017.04.331

1. Introduction

The massive amounts of data being generated in relation to the built environment represent both an opportunity and a challenge for government agencies. Government agencies can derive substantial value from the amounts of data they collect; they can enable a user to draw valuable insights, which can enhance services, increase operational efficiency, and increase transparency [1]. Government agencies need to make significant decisions regularly, therefore an intuitive, evidence-based approach for optimisation of the decision making processes is necessary. Visualisation of housing stock data is one practical approach to help in this process [2]. Data integration from multiple sources is a fundamental step to provide a holistic visualisation of existing data to help agencies better understand the existing housing stock, and improve access to information to support evidence based policy.

The need for the development of a Housing Stock Database for New South Wales (NSW) has been identified as a key priority for the NSW Office of Environment and Heritage (OEH). The objective of the database is to hold fundamental information about the characteristics of the existing residential building stock and become a comprehensive repository of all energy and water efficiency assessment programs in NSW. OEH partnered with the SMART Infrastructure Facility (SMART) and the Sustainable Building Research Centre (SBRC), at the University of Wollongong (UOW), to develop a back-end integrated Housing Stock Database and a front-end online and interactive Housing Stock Mapping Dashboard (HSMD).

The database assembles information collected from multiple assessment programs completed by various agencies over the last ten years and consolidate them at Statistical Area Level 1 (SA1) and Local Government Area (LGA) levels. The pilot phase presented in this paper focused on the Illawarra region and will be expanded to NSW state-wide. This project built upon previous work undertaken by SMART's researchers, like the SMART Infrastructure Dashboard [3] and the Energy Efficiency Dashboard [4]

This paper presents the methodology used to develop the HSMD, from data profiling and categorisation to data warehousing, reporting and visualisation, elaborating on specific challenges associated with multiple sources, typologies and timelines. This is followed by a demonstration of the reporting and visualization platform. Conclusions from the pilot phase and suggestions for future research are then presented in the final section.

2. Methodology and approach

The project was conducted in a number of stages. The initial stage of the project involved the identification and sourcing of housing stock related data from various sources in the government and non-government sectors. This process of identifying and negotiating access to data sources was conducted in close collaboration with government partners. The collected data was then analysed and evaluated to assess priority information, and the quality of the data. A data model was designed to organise and standardise the collected datasets and support the development of the HSMD through a central database structure. The steps and approach that was followed towards the development of this platform is described in detail in the following sections.

2.1. Data sourcing

To map and characterise the existing residential building stock of NSW, a significant number of data sources were required. The identification and acquisition of suitable datasets was an essential first step for this project. Thirteen datasets were initially identified by OEH for potential inclusion within the Illawarra pilot database. The identification of additional data sources was an iterative process, which involved close collaboration between the UOW project team, OEH, and the NSW Department of Planning and Environment (DPE). The UOW project team was responsible for data profiling and evaluation upon delivery of the dataset, with input from relevant stakeholders. Items were identified to be evaluated from previous experience, through existing networks, from literature review and from web searches. As a starting point, metadata or a sample of the data was requested to assess the usefulness of the data source. Advice was also sought from the data managers on any reliability and quality issues for each data source. For government managed datasets the request for full access was typically sent through OEH, quoting the NSW Government Open Data Framework [5]. For publicly available or privately managed data sources, access may have been negotiated by any of the project partners. Delivery of the data was managed by the UOW project team.

2.2. Data profiling and processing

Once a dataset was received by the UOW project team, an extensive evaluation and assessment of the sourced data was undertaken. A key objective of this project was to provide as much detail as possible on the characteristics of the NSW housing stock. However, there is an overwhelming number of building related attributes that could conceivably be captured in a database, and therefore a prioritisation of the building attributes was required. A list was created of parameters which are important for understanding the energy performance of a building, regardless of whether information for the specific parameters was contained in any potential and/or sourced datasets. This list of highly important building parameters was developed from the expert knowledge of the project team, a review of the existing literature, and a survey of other experts in the field. The survey, undertaken as part of a concurrent project to define housing typologies for the NSW stock [6], asked architects, energy efficiency experts, residential energy auditors, and other stakeholders to list the most important parameters required to determine the energy efficiency upgrade potential of an existing dwelling. The results of this survey provided validation of the attribute prioritisation.

A second list was created based on the information contained within the available datasets. As each dataset had collected information for a different purpose using a unique survey tool, the specific entries in each dataset did not directly map across databases. Each specific entry in the list was therefore grouped according to the building parameter to which it related; for instance, the entry “Insulation already installed – wall (Y/N)” was grouped to the parameter “Insulation location”. The parameters were then assigned a relative importance in determined the energy performance of a building, using the same methods as above. This parameter prioritisation was an iterative process which involved close collaboration with OEH and DPE, as the relative importance of parameters is dependent upon the purpose for which they will be used. Gaps between important attributes, which are identified and data covered, are analysed through a comparison of this list with the list of highly important parameters identified. These gaps were used to inform the iterative data sourcing process described above.

Once all the accessible datasets were collected and delivered to the project team, a final iteration of the grouping and prioritisation process was undertaken. All specific data was again grouped into general building parameters which were assigned an indicator of importance. A list of the twenty most important attributes (regardless of data coverage), and another list of the twenty most important attributes for which there was good data coverage were prepared. The UOW project team, in collaboration with OEH and DPE, then examined each attribute on these lists, with reference to the original database records. From this process a final list of attributes for inclusion, and a common definition for each attribute, were agreed upon. The project team worked with stakeholders on developing common definitions for each attribute, which involved determining applicable levels of categorisation and aggregation. Some illustrative examples of this process are presented in the following section.

2.2.1. Attribute categorization and aggregation

In order to map various types of energy appliances, the project team identified four datasets that provided information about the energy appliances used in various types of dwellings. Figure 1 shows a meta-level representation of the categories related to this attribute across the four identified datasets. In Dataset (3), each dwelling may contain a number of energy appliances. For example, a dwelling may have a "Small electric heater (about 1kW)" and a "Split system a/c (sized for smaller room e.g. bedroom)". When there are multiple records for energy appliances, there was no simple method to differentiate between main and secondary systems. As a consequence, area-based aggregation (SA1 for example) was achieved through a count of all listed appliances in each property (e.g. an area may have 100 properties but 125 heater types). Further, Dataset (1) contained information about energy appliances that can relate to the overall building itself, as well as appliances found in each unit or apartment. For example, a building may have central hot water and central heating systems, as well as various individual air cooling or secondary heating systems. In this case, the centralised database reported all these systems at the individual unit or apartment level.

Dataset (1) also allocated two living areas for each dwelling (the living room and the bedroom). This reporting structure amplifies the risks of duplication or misinterpretation during an area-based aggregation process (SA1 for example). Again, there was no simple method to differentiate between main and secondary systems. These

inconsistencies around the definition of specific energy appliances across datasets, and issues associated with multi-dwelling buildings added complexity to the attempt to create a unified typology of energy appliances. The current meta-categorisation (Cooling, Heating and Hot Water systems, see Figure 1) may require further refinement with relevant data providers in order to achieve a robust and comprehensive typology of energy appliances.

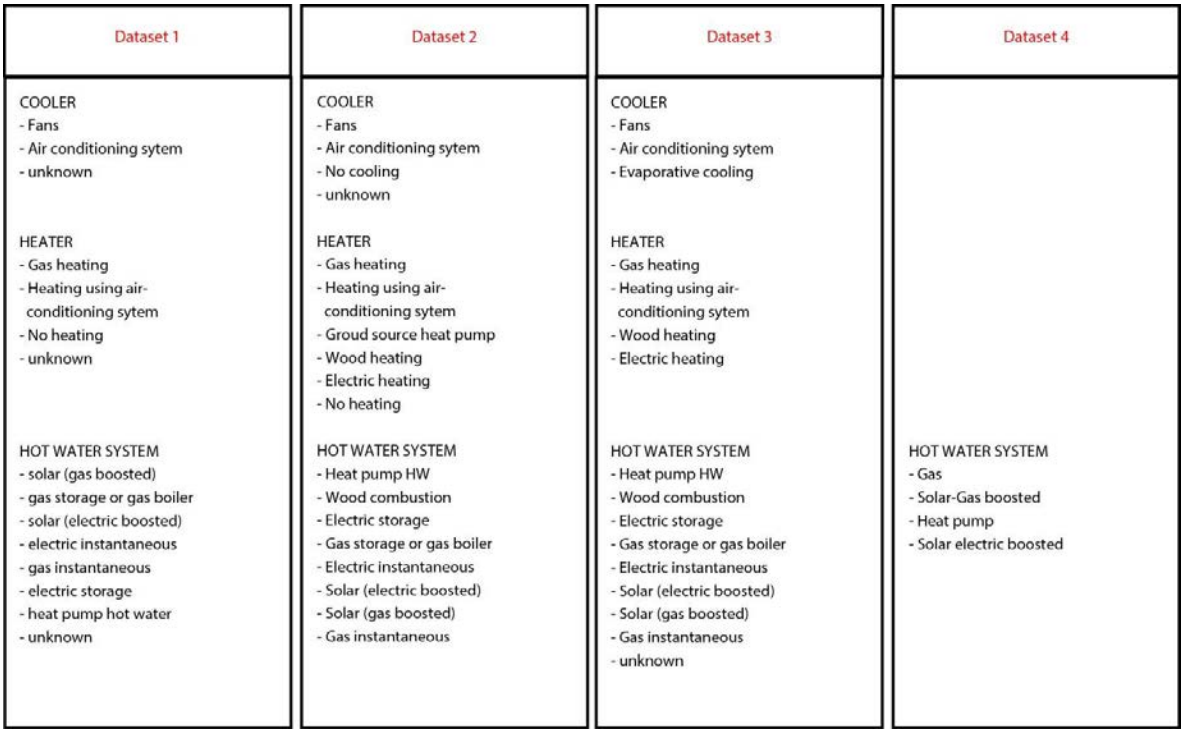

Fig 1. Energy Appliance categories across identified Datasets.

Similar issues were experienced with the dwelling type attribute. The project team identified five datasets that contained information applicable to the dwelling type attribute, with no consistent definition for dwelling structures across the five datasets. Developing a standardised definition for dwelling types across the different datasets presented a major challenge for this project. A first attempt was made to create meta-categories consistent with the Australian Bureau of Statistics (ABS) categories (House, Semi and Unit), however further work with data providers will be required to achieve a robust and comprehensive categorization.

A related issue was determining the orientation of the façade with the most glazing, a proxy measure for whether the building was oriented correctly to optimise solar inputs. The orientation category was reported in three datasets. In one of the datasets, the assessor answered two questions “Are there any glass doors or windows which face south? (Yes, No, Don’t Know)”, and “Are there any glass doors or windows which face North? (Yes, No, Don’t Know)”. These questions do not provide sufficient information to determine the orientation of the façade with the most glazing. However, other datasets provide more details, including orientation and area of glass in each wall. To find out the orientation of the façade with the most glazing, the glazing area for each orientation was calculated, with the largest one considered as the primary orientation.

2.3. Data modelling

Developing a central data repository of the residential building stock of NSW was challenging due to the disparity of the datasets and diversity of the stakeholders. The purpose of the data modelling stage was to build a central database through integrating the disparate data sources which were prioritised and deemed fit for inclusion

from the data profiling process.

2.3.1. Unifying property addresses

An important initial step in the data modelling process was to develop an approach to unify the format of each property address following a standard. The format of a property's address was inconsistent across the datasets due to the use of free text. For example, the property address in one dataset was formatted in six fields. The first four fields were free text format (i.e. the data is entered manually), and the last two fields were predefined (See *Table 1*). The property address in Dataset (2) was defined as three fields. The first two fields were free text format and the last field was predefined (See Table 2).

Table 1. Property address format in Dataset (1).

No	Field	Type of field	Values
1	Street No	Free text	Street number
2	Address1	Free text	Street address
3	Address2	Free text	Street address
4	Suburb	Free text	Suburb
5	Postcode	Dependent	Postcode
6	Council Name	Drop-down menu	Council name

Table 2. Property address format in Dataset (2).

No	Field	Type of field	Values
1	Property Address	Free Text	Property street number and street address
2	Property Suburb	Free Text	Property suburb name
3	Property Postcode	Dependent	Property postcode

To reduce the level of misrepresentation of properties addresses, there was a need to unify the property address format across all datasets. A reference table"property address" was created which contained unique address of each property. The "property_full_address" field in that table was formatted according to the NSW Addressing User Manual. The unique reference "property_id" was created as a unique identifier to link all datasets which contained a property's information (See Figure 2). A geocoding process was then implemented to convert each property address in the database to geographic coordinates.

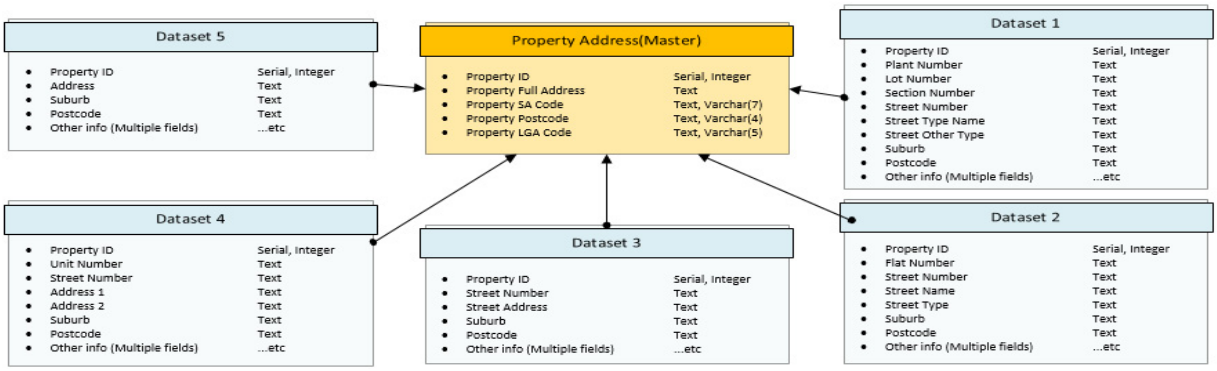

Fig 2. Portion of Data Model used to produce a Unique Property Identifier.

2.4. Database development

A central housing stock database was designed, developed and implemented for scalability, with consideration of the possibility for future expansion of data coverage to include the rest of NSW or Australia.

2.4.1. Selected software and hardware techniques

Virtual machines were used to provide a fast, stable and standards compliant database, with geospatial extensions, a key requirement for analysis and display of residential housing stock data. Visualisation tools were used to generate highly interactive map-based reports that can be used as stand-alone reports or as embedded reports in dashboards. The visualisation tools support Web Mapping Services (WMS) to enrich map-based reports. Data extract, load, and transformation (ETL) scripts were created to implement the ETL functions.

2.4.2. Software components and interactions

The back-end, referred to as the Data Staging Area (DSA), consists of source data (such as MS Excel, flat files) and a staging relational database management system (RDBMS) area to hold data tables prior to a transfer into the main database. This set up can be deployed on a developer's machine, and then migrated to a dedicated virtual machine for the purposes of repeated uploads of data of the same format. The dashboard web-server provides access to the BI layer where reporting, data analysis and visualisation capabilities.

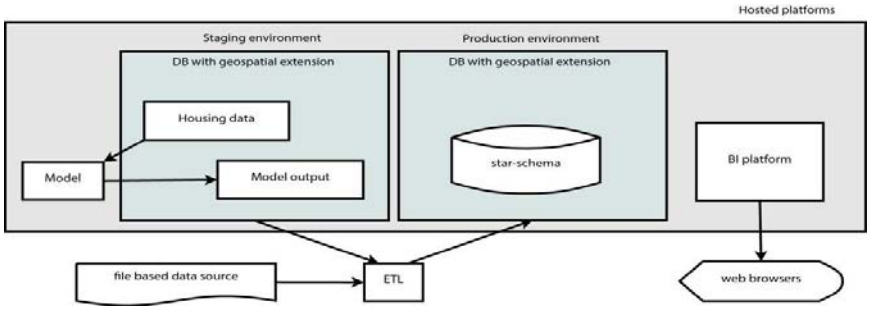

Fig 3. Software components and interactions.

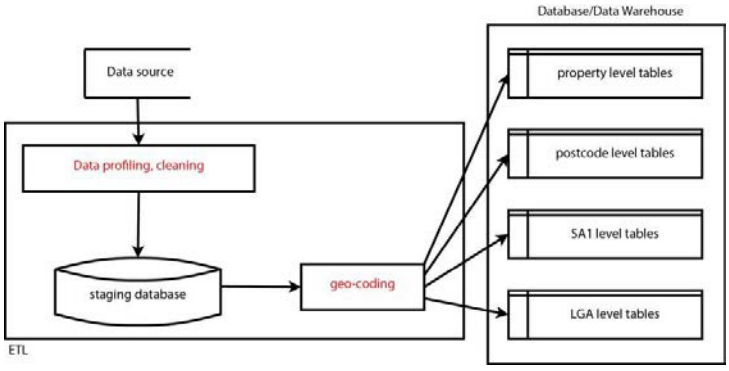

Fig. 4. Example of ETL Data Process.

2.4.3. Extract, Transform, Load (ETL) processing

The ETL processing system (See Figure 4) was designed to meet the requirements of data sourcing, storage and

future expansion of the database. Specifically, it was designed to implement the processes including:

- Download raw data from related websites and sources
- Extract data from the datasets provided by data providers
- Conduct data cleansing, conversion, exception handling and auditing tasks
- Load data to the database
- Monitor and schedule ETL tasks

The details of these main processes are as follows:

1) **Download raw data from related websites and sources.** The project needed to integrate data from related agencies, websites, and databases on a regular basis. The ABS and Bureau of Meteorology (BOM) were the two main websites accessed for this project. From ABS, the project downloads census data and irregular regional statistics. From BOM, the project downloads daily weather (rainfall and temperature) data. To implement this process, several initial Unix Shell (Bash) scripts were developed to download these data automatically.
2) **Extract data in the datasets provided by data providers.** Datasets used in this project were provided by government agecies and private providers. Given that the data from each provider was varied in terms of data format, automatic data transfers were dealt with on a case-by-case basis. The data sourced was commonly stored in spreadsheet files (MS Excel) or .csv (comma separated value). The ETL system was designed to extract relevant data from these files. However, as processing most datasets was one-off task, manual ETL procedure was used for most of datasets.
3) **Conduct data cleansing, convertion, exception handling and auditing tasks.** The database that the ETL system supports was designed based on both the currently accesible data for the project, and envisaged future data acquisitions. Given that more datasets (in addition to those currently at hand) are expected, the database was designed to allow datasets to be effectively added post initial implementation (i.e. tables for future projects). Moreover, new and/or updates to existing data cleansing, converting, exception handling were also developed accordingly.
4) **Loading data to database.** The ETL system transforms appropriate data to corresponding tables in the database from the data staging area.
5) **Monitoring and scheduling ETL tasks.** The ETL system monitors the tasks of processing the data in order to support data verification and exception handling.

2.4.4. Database structure and data model

The database structure is shown in Figure 5. It is composed of two main components: storage and reporting.
The storage component contains seven table “schemas”:

- Referrring schema: contains shared reference information to other schemas. The shared reference includes attributes such as calendar, address naming standard, and BOM weather station list.
- Staging schema: contains cleaned raw data.
- Demography schema: contains demographic, household, and dwelling features at SA1, POA, and LGA levels from ABS.
- Housing stock schema: contains housing stock, building structure and relevant features.
- Government programs schema: contains NSW government relevant programs.
- Non-government programs schema: contains non-government relevant programs or government programs that are managed and implemented by non-government organisations.
- Geospatial schema: contains geographic hierarchy and polygons from ABS.

The reporting component contains tables used for reporting purposes at SA1, POA (Postal Area), and LGA levels. The reporting tables were used to support visualisation.

This project developed a scalable data model. The data model contains three main components: a central address

repository (Property Address), a data storage layer and a data display layer, as shown in Figure 6. The central address repository collected detailed address and geo-location of each individual property or building. It provided unique reference of an individual property or building to data storage and data display layers.

The scalable data storage layer holds both publically available, and private and/or internal-use only data. Depending on its granularity, a dataset can be linked to the address repository or aggregated and combined at the YellowFin BI layer.

The data display layer maintains aggregated data in topics of interest and at given geographic hierarchies. The aggregated data is accessible through employed BI software or tools. The data display layer is the only interface to access data, which prevents the access to private data at data storage layer. Both the data storage layer and the data display layer can be scaled to include data from public services and private sector, further themes and areas.

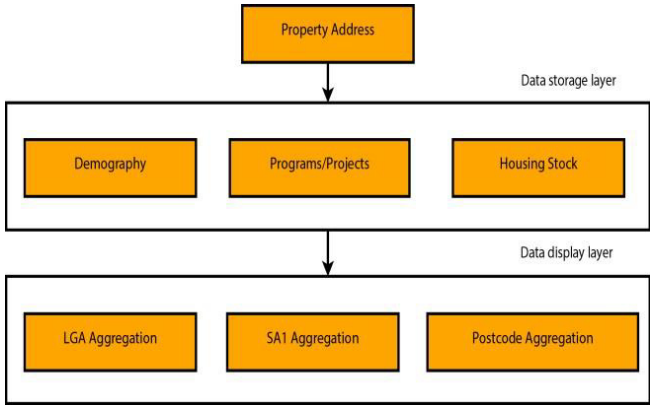

Fig 5: Main Structure of Database

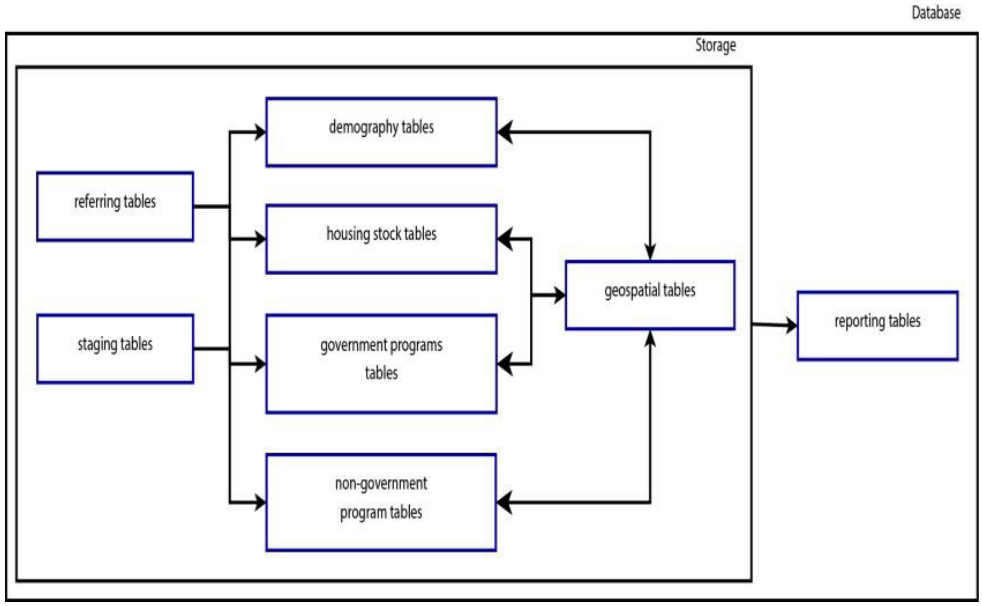

Figure 6 - Data model.

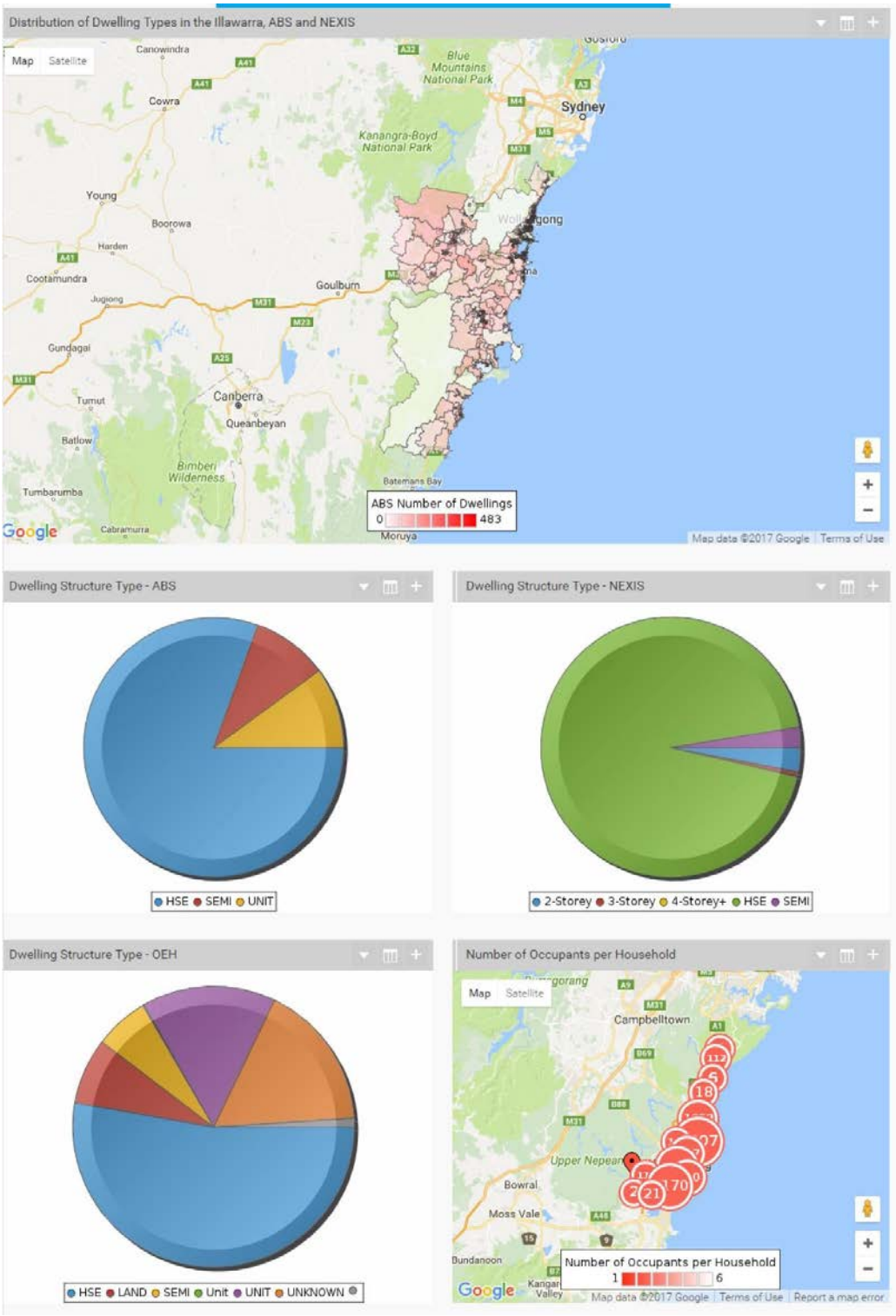

Fig 7: Data model and database illustration.

2.5. Reporting and visualisation development

The reporting and visualisation component of the project was intended to demonstrate the analytical potential of the final database. As part of the BI development process, The project team developed a proof of concept, online analytical and reporting portal, consisting of a number of dashboards, interactive reports and maps, designed to showcase the analystical potential of the underlying House Stock Mapping Database (HSMD). The dashboard was designed to allow users to easily visualise and interpret data within the HSMD. The dashboards displayed the capabilities of the visualisation tool, and provided insights into the included data. Access to the dashboards was initially provided to internal and external stakeholder groups for feedback purposes only. An example of the provided dashboards is shown in Figure 7. The HSMD provided the user with a visually rich interface with easy-to-use controls including filters, pull down menus, and drill through options. Most datasets could be visualised and presented through the platform at a number of geographical levels; i.e. LGA, SA1, and individual property, subject to the level of granularity of the available datasets and privacy constraints.

3. Discussion

Datasets included in the pilot project resulted from a range of previous government initiatives, public surveys, and other projects were sourced from various assessment programs conducted over the last ten years in the Illawarra region, resulting in heterogeneous formats, typologies and timelines. For instance, one dataset contained only data for developments after 2006, whilst data collected for another dataset referred to existing buildings that were assessed between May 2012 and April 2014. Often different databases would have inputs for an identical property at different points in time. In this situation, it was assumed that the most recent record was the correct record for the purpose of aggregation, although all records were retained in the database. For example, a house that had been granted an insulation rebate in August 2008 and was reported in another dataset in 17/8/2012 and 10/5/2013 with information related to insulation. In this situation, it was assumed that the information provided at the latest date (10/5/2013) was the most suitable to use for this property. It should be noted that no validation has been undertaken yet, and for some discrepencies or duplications might still exist in the data warehouse despite all efforts to date.

Through the course of this pilot project a number of potentially valuable datasets were identified, but for which access was not able to be negotiated. There were a variety of reasons for this, including: i) privacy concerns with releasing data at property address level, ii) lack of resources within the organisation willing to provide data to facilitate the data transfer, iii) short timeframe (2 months for data souring in the pilot project), iv) data format issues, v) commercial considerations, and vi) missing metadata. Within the collated data, several important gaps in the data coverage were identified. The significance of these data gaps on the overall quality of the consolidated database were that they had a direct impact on the level of integration that could be attained across multi-source datasets. This in turn compromised the level of granularity that could be achieved in the database, as well as the level at which key attributes could be standardised and classified. Large scale data collection exercises could be designed to fill these gaps, and to provide a means of validating some of the included information. The development of a standardised definition for key attributes across data sources in future project phases would assist towards easy and meaningful integration across data sources.

Several opportunities were identified for further development of the pilot project, including: i) spatial expansion to cover the NSW or Australia, ii) improvement and enrichment of the database, including addressing key data gaps, through either extra data sourcing or extra data collection, iii) the development of sector specific tools and functionalities, and iv) the extension of the tool into comprehensive bottom-up, robust housing stock model i.e. in accordance with the IEA annex 70 activities [7]. In its current format, the centralised database can support a range of applications, for a range of stakeholders in the Illawarra region. Some potential usages envisaged include:

- *Government* - The data centralised in the database can support and inform effective planning; for example, by illustrating areas with higher, or lower densities of particular attributes (e.g. number of insulated houses versus average income in the area), more effectively target the delivery of new government and market driven activities; and provide a benchmark for program evaluation.
- *Householder* - If the database and visualisation were to be made publicly available, a householder would be

able to benchmark economic and sustainability indices of their home against similar homes in their region.

- *Industry and Businesses* – The database can provide quantitative information on current market conditions and market penetration of a range of technologies; it can inform product and service development for relevant businesses and corporations servicing the residential building sector. This information can help provide greater certainty in investment decisions, as well as aid in the development of products and services that most effectively address the shortcomings in the economic, social and environmental performance of NSW dwellings.
- *Energy Utilities* - The tool can help identify potential energy saving opportunities, areas of high penetration of distributed generation technologies (e.g. PV panels), to anticipate demand on the network at a high resolution.
- *Researchers* – The HSMD can form the basis of further study where academics and analysts can interrogate and analyse different datasets. It will provide a more integrated view of operations and markets, as well as broader access to data and can facilitatemulti-disciplinary research.

4. Conclusion

This paper has presented the results of a pilot project to develop a centralised housing stock map database for NSW, Australia. The presented work considered the Illawarra region, and focused on building attributes which will impact the sustainability performance of a dwelling. An extensive data sourcing and processing exercise was undertaken for this project, which uncovered a number of useful datasets for inclusion which had resulted from previous programs. The datasets were collected from government agencies, the UOW data repository, and other sources. Each dataset was then rigorously evaluated, and the attributes contained within categorised and prioritised. The datasets deemed valuable for inclusion were then checked for data quality and formatting issues and a common definition for key attributes was developed. The main challenge was to have a single view of the characteristics of the existing housing stock according to a range of building attributes. To overcome the technical challenge UOW created a centralised database. The database was then used as a single source to visualise the collected datasets through a geo-spatial dashboard. The visualisation provided a flexible way to view the data, and thereby draw insights into the NSW housing stock. The creation of a common definition of the housing attributes, which often differed across data sources, was critical in how the data are presented and evaluated. A number of data issues were addressed, and potential future improvements identified. An important recommendation from this project is that there is a need for a consistent standardised data collection scheme for buildings and their characteristics in order to avoid errors or failed merging attempts when combining data from different datasets.

Acknowledgements

Authors wish to acknowledge NSW Office of Environment and Heritage (OEH) for their funding of this project. They are also thankful to relevant staff from OEH and from NSW Department of Planning & Environment for their support.

References

[1] Kim, G.-H., S. Trimi, and J.-H. Chung, (2014). Big-data applications in the government sector. Communications of the ACM, 57(3): p. 78-85.
[2] Davila, C.C., C. Reinhart, and J. Bemis, Modeling Boston: A workflow for the generation of complete urban building energy demand models from existing urban geospatial datasets.
[3] Wickramasuriya, R., Ma, J., Berryman, M. & Perez, P. (2013). Using geospatial business intelligence to support regional infrastructure governance. Knowledge-Based Systems, 53 80-89.
[4] Safadi, M., Ma, J., Wickramasuriya, R., Somashekar, V., El Fakih, O., Yalan, L., (2013). Energy Efficiency Dashboard for Small Businesses in the Illawarra. International Symposium for Next Generation Infrastructure, Wollongong, Australia.
[5] NSW ICT Strategy, (2013). *NSW Government Open Data Policy*. [ONLINE] Available at: https://www.finance.nsw.gov.au/ict/resources/nsw-government-open-data-policy. [Accessed 18 November 2015].
[6] Daly D, Kokogiannakis G, Aghdaei N, Cooper P, (2016). NSW Housing Typology Development Project: Final Report, Sustainable Buildings Research Centre, Wollongong, NSW, Australia.
[7] RCUK Centre for Energy Epidemiology. 2015. *EBC Annex 70 – Building Energy Epidemiology*. [ONLINE] Available at: https://energyepidemiology.org. [Accessed 18 February 2016].